JURNAL TEKNOINFO
Volume 18, Nomor 2, July 2024, Page 458-475
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

LUNG CANCER DETECTION AND CLASSIFICATION BASED ON DEEP LEARNING: A REVIEW

Hivi Kamal Ismael[1], Adnan Mohsin Abdulazeez[2]

[1] Technical College of Informatic-Akre / Department of Information Technology, Akre University for Applied Science
[2] Technical College of Engineering/Duhok Polytechnic University
[1] Duhok,Iraq
[2] Duhok,Iraq
Email: [1]hevi.kamal@auas.edu.krd, [2]adnan.mohsin@dpu.edu.krd

Abstract

Lung cancer is a significant health problem worldwide because it is difficult to treat and often caused by factors such as smoking and lifestyle choices. Early detection and accurate classification are crucial for assisting patients. Lung cancer remains a major global health challenge due to its late detection and the complexity of its treatment options. Advancements in deep learning, a form of artificial intelligence that mimics the way humans learn, are offering new hopes for earlier detection and more accurate classification of this disease through the analysis of medical images. This review paper explores recent progress in the use of deep learning techniques, specifically focusing on how these methods are applied to improve lung cancer diagnostics. Our study delves into several types of neural networks such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Graph Neural Networks (GNNs), which have been adapted to analyze complex medical imaging data effectively. These networks help in identifying and classifying cancerous tissues from lung scans with a higher degree of accuracy than traditional methods, which rely heavily on human interpretation. We review a variety of models and approaches that demonstrate significant improvements in detecting lung cancer features from imaging studies like CT scans. These models not only enhance the accuracy but also reduce the time needed for diagnosis, which is crucial in improving patient outcomes. The paper discusses the specific roles of these models in automating the detection processes, their efficiency, and how they overcome some of the common challenges in lung cancer diagnosis, such as dealing with ambiguous or incomplete images. Furthermore, we address the challenges still facing deep learning applications in this field, including the need for large, annotated datasets and the computational demands of training complex models. Despite these challenges, the future looks promising due to the continuous improvements in computational power and the increasing availability of medical data.

Keyword: *Lung cancer, classification, deep learning, tumor detection, medical imaging*.

1. INTRODUCING

Cancer is a major problem worldwide, largely due to unhealthy eating habits and lifestyle changes. Recent data indicates that there are approximately 19 million new cancer cases each year, resulting in over 10 million deaths [1]. Lung cancer is a serious health issue because it leads to a lot of deaths. It can be caused by several things, including smoking, drinking too much alcohol, and not eating well. Recently, more young people have started smoking, and this has led to more cases of lung cancer [2]. The imperative for early cancer detection is evident, driving advancements in medical imaging and computational techniques to facilitate diagnosis. Image processing and computer vision technologies have emerged as indispensable assets in the medical domain, offering automated processes that augment the accuracy of disease detection and classification [3-6].

These technologies offer a promising avenue for accurate disease diagnosis through automated processes, providing invaluable support to medical professionals and facilitating confirmatory second opinions [7,8-11]. Traditional cancer diagnosis heavily relies on manual assessment of patient symptoms and medical reports, presenting inherent challenges in accurately assessing tumor characteristics. The intricate nature of these challenges, particularly in the context of lung cancer, necessitates the adoption of more sophisticated analytical approaches [12,13]. Fortunately, the advent of computer vision techniques has revolutionized medical image analysis, equipping clinicians with more efficient and precise diagnostic tools [8]. Notably, recent advancements such as the automated multimodality attention-

JURNAL TEKNOINFO
Volume 18, Nomor 2, July 2024, Page 458-475
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

guided technique proposed by Cheng et al. in 2021 have significantly enhanced tumor detection and classification processes [10].

In recent years, the integration of deep learning (DL) methodologies has propelled advancements in the field of lung cancer diagnosis. Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Deep Reinforcement Learning (DRL), and Graph Neural Networks (GNNs) have emerged as particularly promising avenues for enhancing diagnostic accuracy and efficiency. This paper will thoroughly discuss how these deep learning methods are used for detecting and classifying lung cancer, discussing how well they work, what challenges they face, and what could happen in the future. The goal is to use these advanced methods to improve lung cancer diagnosis and ultimately help patients get better treatment.

The rest paper is structured as follows: Background theory is given in Section 2, giving the background knowledge and context. Section 3 gives a critical review of the literature, summarizing key studies and progress in the field. Section 4 discusses and contrass the various methodologies, together with their strengths and limitations. Finally, the paper concludes with a summarty of the findings in Section 5. The paper reviews deep learning methods like CNNs, RNNs and GNNs for improving lung cancer detection and classification, aiming to enhance diagnostic accuracy and efficiency, and suggests future research directions to improve patient outcomes.

2. BACKGROUND THEORY

2.1 Deep Learning

Deep learning is a component of machine learning that mimics the behavior of the human brain in that it is built on several learning levels and error reduction in intermediate layers [1]. The main feature of deep learning is its ability to work well with features, though it can also acquire knowledge from data [2].Thus, in order to grasp intricate characteristics, deep learning combines the basic features it has acquired through data analysis [3]. Deep learning utilizes complex artificial neural networks with multiple layers, including the Deep Neural Network (DNN), Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN), as depicted in Figure 1. Inspired by the information processing mechanisms observed in the human brain, deep learning operates without the need for predefined rules, relying instead on extensive datasets to correlate inputs with corresponding labels. It is structured through layers of algorithms (referred to as artificial neural networks or ANNs), each offering unique insights into the processed data [11].

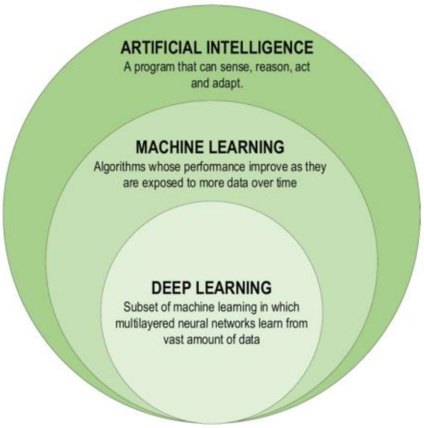

Figure1. Deep learning family [11]

2.2 Classification of DL Approaches

DL methods are divided into three primary groups: unsupervised, partially supervised (also called semi-supervised), and supervised. Additionally, deep reinforcement learning (DRL), also referred to as RL, represents another form of learning technique, primarily categorized as partially supervised (and sometimes unsupervised) learning approaches. Below are the types of DL methods:

a. Deep Supervised Learning

This method pertains to labeled data. When contemplating this approach, the surroundings consist of a set of inputs and their corresponding outputs $(x_t, y_t) \sim \rho$. For example, the intelligent agent $y^{\hat{}}t = f(x_t)$ guesses if the input is x_t and will obtain $i(y^{\hat{}}t, y_t)$ as a loss value. Next, the agent iteratively adjusts the network parameters to refine its estimation of the desired outputs. Once training yields positive results, the agent gains proficiency in resolving queries within its environment. Deep learning offers various supervised learning methods, including recurrent neural networks (RNNs), convolutional neural networks (CNNs), and deep neural networks (DNNs). Within the RNN category, there are gated recurrent units (GRUs) and long short-term memory (LSTM) approaches. A significant advantage of this method is its capacity to utilize prior knowledge for data collection or generation.

b. Deep Semi-Supervised Learning

This approach relies on utilizing partially labeled datasets for learning. At times, it incorporates generative adversarial networks (GANs) and deep reinforcement learning (DRL) similarly. Additionally, recurrent neural networks (RNNs), such as GRUs and LSTMs, are utilized for partially supervised learning. One benefit of this method is its capacity to reduce the dependency on fully labeled data. Conversely, a drawback lies in the possibility of incorrect decisions due to irrelevant input features in the training data. Text document classification stands out as a prime example of applying semi-supervised learning, primarily due to the challenge of acquiring extensive labeled text documents, making semi-supervised learning particularly suited for this task.

c. Deep Unsupervised Learning

This method enables the execution of the learning process even without labeled data, meaning there's no need for predefined categories. Instead, the system learns essential features or internal representations crucial for identifying underlying patterns or connections within the input data. Unsupervised learning encompasses various techniques like generative networks, dimensionality reduction, and clustering. Many deep learning models, such as restricted Boltzmann machines, auto-encoders, and GANs, have shown proficiency in nonlinear dimensionality reduction and clustering tasks. Additionally, recurrent neural networks (RNNs), including GRUs and LSTM, have found applications in unsupervised learning across diverse domains[12].

d. Deep Reinforcement Learning

Reinforcement Learning engages with the environment through interaction, unlike supervised learning, which relies on predefined sample data. It originated in 2013 through Google DeepMind's efforts, spawning various advanced techniques reliant on this approach. For instance, when the agent encounters input environment samples denoted as $x_t \sim \rho$, it predicts $y\hat{}_t = x(f_t)$ and the agent received cost is $c_t \sim P(c_t \mid x_t, y\hat{}_t)$, P here is the unknown probability distribution, then the environment asks a question to the agent. The response provided is a noisy evaluation, often termed as semi-supervised learning. From this principle, various supervised and unsupervised methods emerged. Contrasted with conventional supervised methods, executing this form of learning proves significantly more challenging due to the absence of a straightforward loss function in reinforcement learning. Moreover, two fundamental disparities exist between supervised and reinforcement learning: firstly, there's no complete accessibility to the function necessitating optimization, requiring interaction-based querying; secondly, the state interacted with is grounded in an environment where the input x_t stems from preceding actions [11].

2.3 Deep Learning Approaches

Deep learning techniques utilize artificial neural networks (ANNs) featuring multiple strata of processing units to acquire insights into data patterns. These approaches excel at assimilating intricate representations within expansive datasets, conferring a distinctive edge over traditional machine learning (ML) methodologies. As a result, contemporary cutting-edge approaches to gene expression analysis capitalize on the distinctive competencies offered by these techniques [4] Prevalent neural network architectures encompass fully connected networks (multi-layer perceptron NN), convolutional networks (CNN), recurrent networks (RNN), and graph networks (GNN) [5].

JURNAL TEKNOINFO
Volume 18, Nomor 2, July 2024, Page 458-475
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

a. Multi-Layer Perceptron (MLP)

MLP stands out as a significant form of feedforward neural network used in tasks like pattern recognition, classification, and prediction, primarily employed to address supervised learning tasks [6] MLPs, functioning by linking input to output in a one-way data and calculation process, generally consist of three layers: one for input, one for output, and at least one hidden layer in between [4] These layers are fully connected, with the input layer

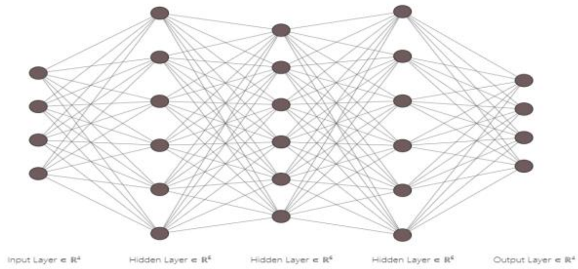

Figure 2.multi-layer perceptron structure

receiving input from the outside world, intermediate layers perform mathematical calculations from input to output, and the final layer generates definitive forecasts. Within each layer, there are nodes or neurons, and the Multilayer Perceptron (MLP) process comprises four main stages. Initially, data is transmitted from the input layer to the output layer. Then, the MLP adjusts its weights between neurons, utilizing a backpropagation algorithm to refine its predictions after handling data for each node [5]. Thirdly, errors are determined by comparing predicted and actual classes, using supervised learning to reduce these errors. Finally, these procedures are repeated through several cycles to enhance and optimize the weights in the learning process. The structure of the MLP is illustrated in Figure 2.

b. Recurrent Neural Networks (RNN)

RNN, encompassing Feedforward Neural Networks, can transmit data across various time steps, as illustrated in Figure 3. Unlike feedforward propagation, which allows information to flow in a singular direction, RNN employs recursion, creating a loop of information as depicted in Figure 3. This recursive approach involves scanning the entire data from left to right, with shared parameters for each time step [13]. Despite its merits, RNN has a limitation – it relies solely on information preceding a point in a sequence for predictions, neglecting any information occurring later in the sequence.

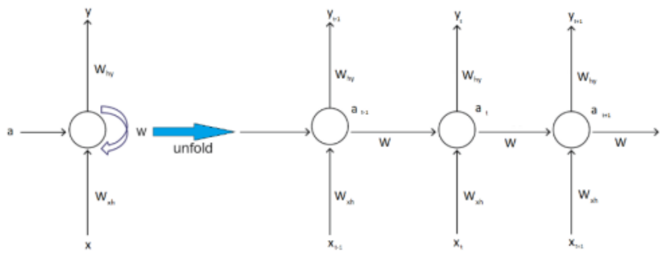

Figure 3. Recurrent Neural Network structure.

c. Convolutional Neural Networks (CNN)

Inspired by the visual processing in animals' brains, CNN is a sophisticated multi-layer neural network pioneered by LeCun et al. Its primary application domains encompass image processing and character recognition, as noted by [14]. The architectural framework involves the initial layer discerning features, followed by intermediate layers that amalgamate these features to generate high-level input characteristics, culminating in a classification process. The accumulated characteristics undergo pooling to reduce dimensionality, and subsequent steps involve convolution and pooling, ultimately feeding into a fully connected multi-layer perceptron [7]. The final layer, the output layer, employs back-propagation techniques to recognize the image's distinctive features, as elucidated by [14]. CNN stands out due to its distinctive attributes, such as local connection and shared weights, contributing to heightened system accuracy and performance. It surpasses other deep learning techniques and stands as the most employed architecture. For a visual

JURNAL TEKNOINFO
Volume 18, Nomor 2, July 2024, Page 458-475
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

representation, refer to Figure 4.illustrating the structure of a convolutional neural network [1].

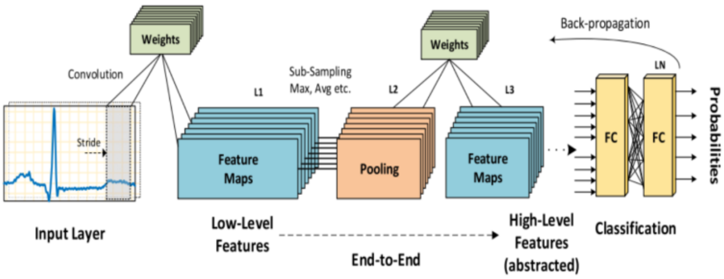

Figure 4. Convalution Neural Network structure

d. Long Short-Term Memory (LSTM)

LSTM networks, falling under the umbrella of recurrent neural networks (RNNs), exhibit a noteworthy proficiency in grasping long-term dependencies, as exemplified in [22]. The architecture of an LSTM involves the intricate construction of a memory cell utilizing logistic and linear units with multiplicative interactions. This design facilitates a dynamic flow of information within the cell: information is admitted through the input gate, expelled when the forget gate is inactive, and accessed for reading by activating the output gate. Such nuanced operations empower LSTMs to effectively capture and retain information over extended sequences, showcasing their prowess in addressing scenarios with prolonged dependencies [4]. Figure 5.illustrating the structure of a Long short-term memory.

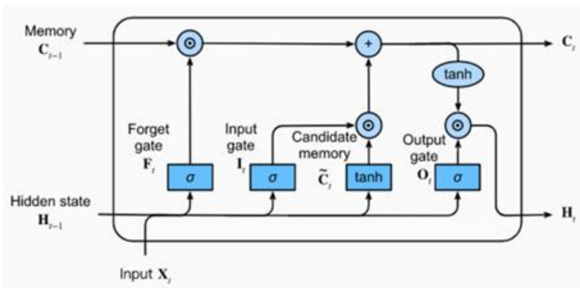

Figure 5. Long short-term memory (LSTM) structure

e. Graph Neural Networks (GNN)

GNNs belong to the realm of deep learning algorithms tailored for the examination and interpretation of structured data encapsulated within graphs. Graphs, comprising interconnected nodes and edges, serve as versatile models to depict relationships and dynamics across diverse domains like social networks, biological systems, citation networks, and recommendation frameworks [8]. The primary objective of GNNs is to acquire nuanced representations of individual nodes within a graph [9]). This entails capturing not only the characteristics of a node's immediate surroundings but also discerning patterns in the broader structural context of the entire graph.

In the realm of GNNs, the process begins with the representation of a graph where nodes symbolize entities, and edges signify relationships between these entities. Each node is endowed with features, offering insights into the corresponding entity. The journey continues with node embeddings, where initial embeddings are assigned to nodes based on their features. GNNs engage in iterative message passing steps, allowing nodes to gather information from their neighbors and update their embeddings accordingly; this involves the exchange of neighborhood information and using learnable aggregation functions such as mean, sum, or attention mechanisms [8]. Stacked aggregation layers further refine node embeddings by assimilating information from increasingly expansive neighborhoods to encapsulate

JURNAL TEKNOINFO
Volume 18, Nomor 2, July 2024, Page 458-475
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

local and global graph structures; some employ graph pooling layers, contributing to hierarchical representation [9]. Ultimately, the process culminates in the output layer, where the final node embeddings derived from these intricate steps can be applied to diverse tasks such as node classification, link prediction, or graph classification.

2.4 Lung Cancer

Carcinogenesis is the unchecked proliferation of one or more cell types. Good tissues do not support the growth of normal cells, and when they do, they separate quickly and become tumors [3]. Primary lung cancer originates elsewhere in the body and spreads to the lungs, while secondary lung cancer starts elsewhere in the body and then spreads from there. It's one of the most aggressive types of cancer and a life-threatening threat to the human body . If this unchecked development can be identified correctly at an early point, it can help to diagnose the likelihood of unnecessary surgery and improve the chance of recovery. Chronic Obstructive Pulmonary (COPD) illness attacks the areas of the lungs and causes diseases such as measles, influenza, pneumonia, and other respiratory issues such as asthma. Small Cell Lung Cancer (SCLC) or oat cell cancer and Non-Small Cell Lung Cancer (NSCLC) are the two main forms of lung cancer that develop and expand in separate ways and may be handled accordingly. Within the non-small cell lung cancer category, there are three subtypes (adenocarcinomas, squamous cell carcinomas, large cell carcinomas) figure 6.show the two types of lung cancer. So Mixed small cell/large cell cancer is a disease that occurs where a patient shows symptoms of both types of cancer. (NSCLC) Adenocarcinoma is more common and progresses more slowly than small cell lung cancer. Small cell lung cancer is linked to smoking which progresses more rapidly by becoming a large tumor that will spread across the body [10].

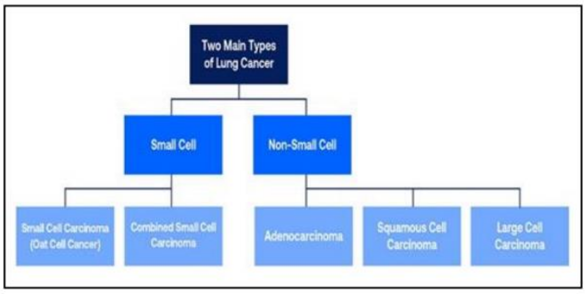

Figure 6. Lung Cancer Types

2.5 Classification

Classification is a method of learning where a program is trained to distinguish between multiple classes using input data. During testing, the classifier predicts the class of new or unknown data samples. This approach proves valuable in medical contexts, particularly in the analysis of lung cancer nodules. It aids in discerning between nodules and non-nodules within the lungs, and further categorizing nodules as benign or malignant. These outcomes empower radiologists and pulmonologists to identify nodules with increased confidence [15]. Classification algorithms can generally be divided into two types: traditional machine learning, which relies on pre-defined features derived from input data, and deep learning, which learns features directly from the input. Deep learning, currently in vogue, excels particularly with large datasets, demonstrating enhanced performance as data size increases. In contrast, traditional algorithms plateau in performance beyond a certain point, lacking the capacity to leverage additional data for improvement [16]. Various algorithms within both categories have been applied across diverse fields such as medical imaging, emotion recognition, and speech processing, yielding promising results. Figure 7. illustrates an intricately crafted deep learning framework designed to distinguish between three specific types of lung cancer (LC) using CT scan images, while also encompassing instances of normal patient scans [17].

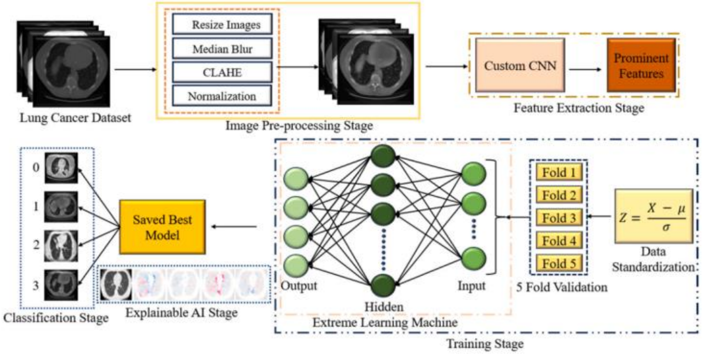

Figure 7. Proposed framework for classification of lung cancer types from CT scan images.

2.6 Metrics for performance evaluation

a. Accuracy

Accuracy (AC) is a metric employed to assess the efficacy of a classification model [18]. In machine learning, especially in tasks involving classification, accuracy indicates the model's proficiency in correctly forecasting the labels of dataset instances. The mathematical expression for calculating accuracy is as follows [19]:

$$Accuracy\ = \frac{(\boldsymbol{TN} + \boldsymbol{TP})}{(\boldsymbol{TN} + \boldsymbol{TP} + \boldsymbol{FN} + \boldsymbol{FP})}$$

Where, TP = True Positives, TN = False Negatives, FP = False Positives, FN = False Negatives.

b. Sensitivity

Sensitivity is the proportion of actual positives that the classifier correctly identifies during the testing process, and can be calculated with respect to the equation [19].

$$Sensitivity\ = \frac{\boldsymbol{TP}}{\boldsymbol{TP} + \boldsymbol{FN}}$$

c. Specificity

Specificity is the proportion of actual negatives that a classifier correctly identifies during the testing phase and can be computed using the following formula [20].

$$Specificity\ = \frac{\boldsymbol{TN}}{\boldsymbol{TN} + \boldsymbol{FP}}$$

And there are some other performance metrics, as follows [13]:

$$Recall\ = \frac{\boldsymbol{TP}}{\boldsymbol{TP} + \boldsymbol{FN}}$$

$$Precision\ = \frac{\boldsymbol{TP}}{\boldsymbol{TP} + \boldsymbol{FP}}$$

$$F1 - score\ = \frac{\boldsymbol{2} * \boldsymbol{Precision} * \boldsymbol{Recall}}{\boldsymbol{Precision} * \boldsymbol{Recall}}$$

3. LITERATURE REVIEW

In recent days, several deep learning techniques were introduced by researchers mainly to improve the accuracy of detecting lung cancer from CT images.

Dodia et. al. (2022) [15] examined deep learning algorithms for detecting lung cancer from CT scans, focusing on methods like CNNs, U-Nets, DenseNets, and their preprocessing and segmentation techniques. It contrasted traditional models like AlexNet, VGG, and ResNet with those tailored for medical imaging, such as 3D convolutional networks. Performance metrics across studies—accuracy, sensitivity, specificity, and AUC—were thoroughly analyzed, revealing the effectiveness of various approaches in lung nodule detection and suggesting high accuracy and sensitivity in some studies. Despite the progress in utilizing deep learning for lung cancer detection, challenges like data imbalance, the need for larger datasets, and computational demands persisted. Future research directions included exploring multimodal data, enhancing model efficiency, and devising solutions for data imbalance.

Cao et. al. (2023) [21] presented E2EFP-MIL, a cutting-edge deep learning method for classifying lung cancer subtypes from WSIs. This end-to-end, weakly supervised approach was data-efficient and used iterative patch sampling, a feature pyramid module, and feature aggregation to extract and identify critical morphological patterns. Trained on 1007 WSIs from the TCGA dataset, E2EFP-MIL achieved high accuracy with AUCs between 0.95 and 0.97. Validation on nearly 1600 external WSIs from the U.S. and China confirmed its effectiveness and data efficiency, requiring only 100-200 training images to achieve an AUC of >0.9. Outperforming other MIL-based methods, E2EFP-MIL's success demonstrated the potential of AI in improving diagnostic accuracy for lung cancer, suggesting significant implications for patient management and treatment.

Nahiduzzaman et. al. (2024) [16] presented a novel approach that combined a lightweight parallel depth-wise separable convolutional neural network (LPDCNN) with a ridge regression extreme learning machine (Ridge-ELM) for classifying three lung cancer types and normal lung tissue using CT images. The methodology enhanced image quality and reduced noise with contrast-limited adaptive histogram equalization (CLAHE) and Gaussian blur. The LPDCNN extracted key features efficiently, while the Ridge-ELM improved classification performance. The framework achieved high average recall and accuracy of 98.25% and 98.40%, respectively, in four-class classifications and 99.70% in binary classifications. It was also highly efficient, with a testing time of 0.003 seconds. Integration of SHAP (Shapley Additive Explanations) provided insights into decision-making, increasing confidence in real-world lung cancer diagnoses.

Shyamala Bharathi and Shalini (2024) [22] highlighted advancements in lung cancer detection using deep learning with CT and PET images. It introduced a model combining Adaptive Dilated Convolution Neural Network (AD-CNN) with a Modified Initial Velocity-based Capuchin Search Algorithm (MIV-CapSA) for optimized image fusion. Using TransUnet3+ for segmentation and Hybrid Attention-based Deep Networks (HADN) for classification, the model outperformed traditional methods in accuracy, sensitivity, specificity, and efficiency. Validation showed its effectiveness in classifying lung cancer types, marking a shift towards more advanced deep learning applications in medical imaging for early and accurate diagnosis.

Gumma et. al. (2022) [23] discussed the application of CNNs in lung cancer detection through CT scans, detailing the process from image pre-processing to the classification of nodules into benign or malignant. It emphasized the evolution from 2D to 3D CNNs, noting that 3D models provided a more comprehensive analysis by incorporating depth, leading to higher accuracy in detection and classification compared to traditional methods. The review highlighted that CNN-based approaches, particularly 3D CNNs, had significantly improved the sensitivity and reduced false positives in lung cancer detection. It concluded that deep learning, with a focus on CNNs, offered great potential for enhancing early and accurate lung cancer diagnostics, marking a pivotal shift towards more advanced and reliable detection models.

Xu et. al. (2022) [24] introduced ISANET, a cutting-edge method that leveraged convolutional neural networks (CNNs) enhanced with channel and spatial attention mechanisms for the classification and detection of non-small cell lung cancer (NSCLC) from chest CT scans. Developed to address the challenge of distinguishing between NSCLC subtypes due to overlapping imaging features, ISANET enhanced the InceptionV3 architecture by incorporating mechanisms that focused on pathological areas, improving differentiation between lung squamous cell carcinoma, lung adenocarcinoma, and normal lung tissues. Comparative evaluations against traditional models on a lung cancer dataset demonstrated ISANET's superior accuracy, achieving 95.24% and 98.14% on two public datasets. Ablation studies highlighted the significant role of the attention mechanisms in achieving these results, indicating ISANET's potential in enhancing CT-based lung cancer diagnostics and early subtype identification through advanced CNNs and attention mechanisms.

Raju and Rao (2022) [25] introduced a deep learning approach for classifying colon and lung cancer using histopathological images. Utilizing MobileNetV2 and InceptionResnetV2 models, enhanced with GradCam and SmoothGrad visualization techniques, the method achieved up to 99.95% accuracy. Tested on a dataset of five types of colon and lung tissues, these models proved effective in accurately identifying cancer tissues. The approach aimed to assist in developing automated systems for early detection of colon and lung cancers, potentially reducing mortality rates.

Kalaivani et. al. (2020) [26] introduced a novel deep learning approach for detecting and classifying lung cancer from CT scans using DenseNet and AdaBoost algorithms. Preprocessing involved a set of 201 lung images, with 85% used for training and 15% for testing, and data augmentation to enhance model robustness. Combining DenseNet's image classification and AdaBoost's focus on challenging instances, the method achieved a classification accuracy of 90.85%. This development marked a significant advance in medical imaging for lung cancer detection, offering promising avenues for future research and practical healthcare applications.

Agarwal et. al. (2021) [27] introduced a novel method using the AlexNet Convolutional Neural Network (CNN) for classifying lung tumors from CT images. By employing transfer learning with AlexNet, this approach significantly outperformed traditional neural network methods in accuracy. The process involved preprocessing CT images, extracting lung regions through multilevel thresholding, and distinguishing between malignant and benign tumors using AlexNet-CNN. After retraining, the model achieved a remarkable 96% accuracy. This development highlighted the potential of deep learning to improve the accuracy and efficiency of lung cancer diagnosis, marking a significant advancement in the early detection of lung cancer through medical imaging analysis.

Asuntha and Srinivasan (2020) [28] introduced an advanced deep learning model, the Fuzzy Particle Swarm Optimization Convolutional Neural Network (FPSOCNN), an advanced deep learning model for detecting and classifying lung cancer nodules from CT scans. The study involved preprocessing, segmentation, feature extraction, and classification using techniques like histogram equalization and Adaptive Bilateral Filtering. FPSOCNN, enhanced with Fuzzy Particle Swarm Optimization for feature selection, outperformed traditional CNN models in accuracy, sensitivity, specificity, and computational efficiency. The research highlighted FPSOCNN's potential to improve early lung cancer diagnosis and management.

Aluka et. al. (2023) [29] evaluated pre-trained deep learning models for lung cancer detection using a 5000-image dataset. It focused on VGG-16, Inception V3, and ResNet50 models, employing transfer learning for improved accuracy. VGG-16, with fine-tuning and image augmentation, significantly outperformed the others, achieving 96% training and 93% validation accuracy. Inception V3 and ResNet50 did not achieve comparable results. The paper concluded that VGG-16 was the most effective model for lung cancer detection, underscoring its potential in medical imaging and contributing to the advancement of accurate and efficient cancer detection methods.

Pradhan et. al. (2023) [30] introduced an innovative lung cancer diagnosis method using patient health records, combining principal component analysis (PCA) and t-SNE for feature extraction. These were enhanced by the SA-SLnO algorithm, which optimized feature weights and RNN neuron counts to minimize Mean Squared Error (MSE) and improve accuracy, sensitivity, and precision. This model outperformed existing methods with its superior solution quality and speed, offering a stable, efficient, and less complex approach for early and accurate lung cancer detection through advanced algorithms.

Khatun et. al. (2023) [31] presented a deep learning strategy for lung cancer diagnosis using histopathology images, focusing on NSCLC subtypes. A dataset of 15,000 images, expanded from 750 originals, was used to evaluate four CNN models: VGG19, ResNet50, EfficientNetB7, and MobileNetV2. ResNet50 was the most effective, achieving 98% accuracy, with VGG19 also performing strongly. The study highlighted ResNet50's potential in enhancing the accuracy and efficiency of lung cancer detection, suggesting a promising avenue for further research and model refinement.

Nafea et. al. (2023) [32] discussed a study that applied the EfficientNet B3 model to identify lung cancer types through CT scans, aiming to enhance detection accuracy and patient care. Using a dataset of 1000 CT scans from Kaggle, the study employed data augmentation and transfer learning from the ImageNet dataset to improve training. The EfficientNet B3, customized with extra classification layers, achieved a 96% accuracy rate, outperforming previous models. This research demonstrated the model’s effectiveness in medical diagnostics and set the stage for further exploration of EfficientNet architectures in early cancer detection.

Shankara et. al. (2023) [33] showcased a novel approach to lung cancer detection using CNNs. Focused on improving diagnosis accuracy to potentially increase survival rates, the model was trained on a substantial dataset from the Lung Image Database Consortium (LIDC), featuring both cancerous and non-cancerous lung images. Employing image preprocessing and watershed segmentation, the researchers enhanced the CT scan images for more precise analysis. The CNN model, used for feature extraction, training, and classification, demonstrated high efficacy by accurately identifying lung CT images as cancerous or normal with an accuracy of 92.96%, a sensitivity of 97.45%, and a specificity of 86.08%. This study illustrated the power of CNNs in automating lung cancer detection, offering a significant step forward for early diagnosis and treatment. It also set the stage for future research into CNN applications for cancer detection, emphasizing lung cancer.

Prasanna et. al. (2023) [34] presented a study on a novel hybrid deep learning approach for early lung cancer detection using advanced CNNs and DNN architectures. Employing models like CNN, SqueezeNet, and MobileNet with a CT scan dataset, the hybrid DNN model achieved a notable accuracy of 95.21%, outperforming existing models. SqueezeNet was particularly effective due to its rapid processing and balanced performance. The research pointed to future exploration of different segmentation methods and 3D volumetric analysis, highlighting the potential of hybrid architectures in enhancing diagnosis accuracy and patient care in medical imaging.

Abid et. al. (2023) [35] discussed a study that used a CNN-based approach to enhance lung cancer detection from CT scans, focusing on early nodule identification with deep learning. Employing the LUNA-16 and LIDC datasets, the study improved image analysis through data augmentation, noise removal, and segmentation. It tested pre-trained models like DenseNet, AlexNet, and VGG-16, with DenseNet emerging as the most effective, achieving 98% accuracy, 98.93% sensitivity, and 99% specificity. This research underscored the effectiveness of CNNs, particularly DenseNet, in early lung cancer detection, contributing significantly to advances in computer-aided diagnosis systems.

Kumar et. al. (2023) [36] highlighted a major advancement in lung cancer diagnosis using an advanced UNet model for early detection. Utilizing deep learning and metaheuristics, the research developed a convolutional neural network with the predator technique, achieving high detection rates (76.51% for 140 keV and 81.58% for 60 keV) and surpassing qualified physicians. The model showed high accuracy (93.4%), sensitivity (98.4%), and specificity (97.1%), with a low error rate (1.6%), outperforming traditional CNN models in lung segmentation. This breakthrough significantly improved lung cancer classification from CT scans, enhancing early diagnosis and patient outcomes.

Biradar et. al. (2022) [37] introduced a 2D CNN model for diagnosing lung cancer using CT scans from the Kaggle and Luna16 datasets. It involved preprocessing scans to reduce noise and converting them into 2D slices, with K-means clustering and SVM for segmenting and classifying nodules. The 2D CNN architecture, including convolutional, max-pooling, and fully connected layers, extracted key features like lung nodule edges. Achieving 88.76% accuracy, the model outperformed K-means, KNN, and SVM methods, proving effective in identifying lung cancer and showing potential as a reliable tool for early diagnosis. The research suggested future exploration into precise nodule detection and measurement improvements.

Anand et. al. (2023) [38] focused on differentiating between lung and colon cancer types using deep convolutional neural networks (CNN). Utilizing a dataset of 25,000 histopathology images, the research developed a CNN model with four convolutional blocks, optimized for batch size, epochs, and optimizers. The model achieved a 99.0% accuracy rate, surpassing traditional methods like K-means, KNN, and SVM, with precision, recall, and F1-scores between 95% and 98%. This research demonstrated the CNN model's robust capability in cancer classification, significantly contributing to early cancer detection and improving diagnostic processes and patient outcomes. It also laid the groundwork for future in-depth classification studies in medical imaging.

Venkatesh and Bojja (2022) [39] presented a new lung cancer detection method using CT images, combining deep learning, cuckoo search optimization, and IoT. It preprocessed images, segmented nodules with Otsu thresholding and cuckoo search, and extracted features using Local Binary Patterns (LBP). CNNs classified nodules, achieving around 98% accuracy and a Mean Square Error of 1.5. IoT integration via a Raspberry Pi allowed efficient data sharing, enhancing clinical use and supporting early diagnosis and treatment planning.

Malligeswari and Kavya (2024) [40] introduced a novel approach for classifying lung cancer into small cell (SCLC) and non-small cell (NSCLC) categories using deep learning and PET/CT scans. Researchers used a dataset from radiogenomics databases, applied Gaussian smoothing for preprocessing, and developed a segmentation model with the Res-U-Net architecture, complemented by a Support Vector Machine (SVM) classifier. The methodology achieved a high accuracy rate of 96.45%, outperforming other neural networks. The study underscored the potential of this method to improve lung cancer classification, suggesting it could aid treatment selection and warranted further clinical validation and exploration for other cancer types.

Pandey and Bhandari (2023) [17] explored modern healthcare strategies for lung cancer detection and classification, focusing on AI integration, especially ML and DL. It emphasized AI's potential to improve accuracy and efficiency in diagnosis, addressing challenges in detecting cancer within lung structures. Through a review of literature from 2012 to 2022, it evaluated various ML and DL models, highlighting the progression towards sophisticated DL models like CNNs. These models showed significant advancements, particularly in lung nodule segmentation and classification, potentially aiding earlier detection and reducing mortality rates. The paper concluded that AI-based models, particularly in ML and DL, represented a crucial advancement in lung cancer diagnosis, urging further research to refine and apply these models in clinical settings for more effective treatment strategies.

Pandian et. al. (2022) [41] developed an automated system for early lung cancer detection using deep learning techniques, specifically CNN and GoogleNet, applied to CT scan images. Addressing the need for accurate early diagnosis due to the high mortality rate of lung cancer, the research proposed an enhanced detection algorithm based on the VGG-16 architecture. It distinguished between various lung cancer types and normal images, utilizing a dataset of

100 images per class, split into training and validation sets. The findings highlighted a significant achievement, with the deep learning models reaching an overall accuracy of 98% in detecting and classifying lung cancer, demonstrating the potential of CNN and GoogleNet in improving diagnostic processes and treatment planning with high efficiency.

Islam et. al. (2023) [42] introduced an innovative approach to detect and classify lung conditions like COVID-19 and lung cancer, leveraging a combination of Deep Convolutional Neural Networks (DCNN) and Gated Recurrent Units (GRU) with the addition of Explainable Artificial Intelligence (XAI) for enhanced interpretability. The method applied XAI techniques such as LIME, SHAP, and Grad-CAM to provide insights into the model's decision-making, thereby increasing its transparency. Tested on distinct datasets for COVID-19 and lung cancer, the model demonstrated exceptional accuracies of 99.30% and 98.97%, respectively, surpassing existing solutions and highlighting its effectiveness and practicality for real-world use. The integration of XAI not only addressed interpretability challenges in deep learning but also aided in making more informed clinical decisions, ultimately benefiting patient care and treatment planning.

W. Abdul (2020) [43] introduced the ALCDC system, an automated tool using a CNN for lung cancer detection from CT images. It emphasized the importance of early detection for improving treatment outcomes and utilized deep learning to enhance accuracy, achieving 97.2%. Validated with LIDC and IDRI images and designed to prevent overfitting with structured layers, the system demonstrated the effectiveness of CNN-based approaches in advancing medical imaging and diagnostics.

Mothkur and Veerappa (2022) [44] explored an innovative deep learning approach to detect lung cancer by analyzing CT images with lightweight deep neural networks (DNNs) designed for low memory use. This research investigated the effectiveness of three DNNs—vanilla 2D CNN, 2D SqueezeNet, and 2D MobileNet—in identifying early-stage lung cancer, addressing the limitations of traditional computer-aided diagnosis systems, such as high false-positive rates. The study found that the lightweight DNN models, particularly SqueezeNet, offered a promising solution with an overall accuracy of 85.21%, demonstrating a balance between specificity and sensitivity. SqueezeNet was highlighted as the superior model due to its quick processing time and balanced performance, suggesting a shift towards simpler models for medical image analysis. The authors proposed future work to further improve model accuracy through advanced segmentation techniques and the application of these models to 3D volumetric slices, aiming to enhance lung cancer detection and classification.

Mathivanan et. al. (2024) [45] described the Multimodal Fusion Deep Neural Network (MFDNN), an innovative AI model that improved lung cancer diagnosis by integrating medical imaging, genomics, and clinical data. MFDNN tackled ethical concerns, AI interpretability, and strict regulatory compliance, outperforming traditional models like CNN, DNN, and ResNet with a 92.5% accuracy rate. It also achieved precision and recall rates of 87.4% and 86.4%, respectively, and an F1-score of 86.2. These results highlighted MFDNN's effectiveness in accurately diagnosing lung cancer, significantly advancing AI's role in enhancing diagnostic accuracy and setting new standards for early detection and classification in healthcare.

Naseer et. al. (2023) [46] presented an automated system for lung cancer detection from CT scans using computational intelligence. This system processed through lobe segmentation, nodule extraction, and classification phases, utilizing modified U-Net and AlexNet architectures alongside SVM. The modified U-Net excelled in segmenting lung lobes and extracting nodules, while the AlexNet-SVM combination classified nodules into cancerous or non-cancerous with high accuracy. Tested on the LUNA16 dataset, the approach achieved a 97.98% accuracy rate, demonstrating its capability for precise lung cancer detection and minimizing misdiagnosis. This method represented a significant advancement in AI-driven lung cancer diagnostics, improving diagnostic accuracy and establishing a new standard in the field.

Malafaia et. al. (2022) [47] investigated the robustness and interpretability of DL models in lung cancer detection using CT images, addressing the challenge of diagnosing from limited datasets. The research employed Explainable AI (XAI) techniques—Saliency Maps, Integrated Gradients, Layer-wise Relevance Propagation, and Deep Learning Important Features—to provide visual explanations of model predictions, emphasizing the necessity of interpretability for sparse training data. The study's findings demonstrated the DL model's stability and consistent focus on clinically relevant areas across different training sets, verified by visual explanations that illuminated diagnostically significant regions within CT scans. Concluding that XAI enhances model transparency and interpretability, the study affirmed the importance of XAI in making DL models in lung cancer classification more understandable and trustworthy, underscoring the critical role of explainable methods in healthcare, where accurate and reliable diagnostics are paramount.

Wankhade and Vigneshwari (2023) [48] introduced a pioneering hybrid deep learning method, CCDC-HNN, for early lung cancer detection using the LIDC-IDRI database for CT scan analysis. It combined 3D-Convolutional Neural Networks (3D-CNN) and Recurrent Neural Networks (RNN) to enhance diagnostic accuracy and distinguish between benign and malignant tumors. The CCDC-HNN method significantly improved over traditional and some

JURNAL TEKNOINFO
Volume 18, Nomor 2, July 2024, Page 458-475
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

current deep learning methods, achieving high accuracy, sensitivity, and specificity in lung cancer stage identification. This research marked a major advancement in lung cancer diagnostics through AI, showcasing the effectiveness of hybrid neural network models for early and accurate detection.

Lee et. al. (2022) [49] presented an effectiveness of various CNN architectures and tile sizes for NSCLC detection in whole slide images (WSIs). The study tested VGG19, InceptionResnetV2, and EfficientNet b3 models across tile sizes from 296 to 10,000 pixels, using 87 annotated WSIs. It found that tile sizes between 500 and 1,000 pixels yielded the best results, with VGG19 slightly outperforming the others at 500x500 pixels. This optimal size balanced the inclusion of essential pathological features with computational efficiency. The findings emphasized that selecting appropriate CNN architectures and tile sizes significantly improved NSCLC detection accuracy in WSIs, highlighting their crucial role in enhancing AI diagnostics in pathology.

Shakeel et. al. (2022) [50] outlined a new method for detecting lung cancer from CT images, employing advanced deep neural networks and ensemble classifiers to address the challenges posed by low-quality images. This approach used a dataset of non-small cell lung cancer CT scans and featured techniques like a multilevel brightness-preserving method for image enhancement, enhanced deep neural network segmentation, and feature extraction. Feature selection was done using a hybrid spiral optimization and intelligent-generalized rough set approach, with classification handled by an ensemble classifier. Evaluated in MATLAB, this method showed significant improvements in accuracy over existing techniques, presenting a promising advancement for early lung cancer detection in clinical settings.

Table 1. Summary of Literature Review

Authors & Reference	Year	Datasets	Algorithms	Key Results	Advantages	Disadvantages
Dodia et. al. [15]	2022	CT scans	CNNs, U-Nets, DenseNets, traditional models	high accuracy and sensitivity in lung nodule detection	Tailored models for medical imaging, thorough analysis of performance metrics	Data imbalance, need for larger datasets, computational demands
Cao et. al. [21]	2023	Whole slide images (WSIs)	E2EFP-MIL	High accuracy (AUCs between 0.95 and 0.97)	Data-efficient approach, high generalizability	Requires high-quality training data
Nahiduzzaman et. al. [16]	2024	CT images	LPDCNN, Ridge-ELM	LPDCNN with Ridge-ELM achieved 98.25% recall and 98.40% accuracy in classifying lung cancer types.	Lightweight model, efficient classification	None specified, but potentially complex implementation
Shyamala Bharathi and Shalini [22]	2024	CT and PET images	AD-CNN, MIV-CapSA, HADN	AD-CNN and MIV-CapSA models outperformed traditional methods in accuracy and efficiency	Enhanced image clarity, detection precision	May require sophisticated hardware for optimal performance
Gumma et. al. [23]	2022	CT scans	CNNs (3D)	3D CNNs achieved 91.23% accuracy with 3.99 false positives per scan.	Comprehensive analysis of CNN evolution for lung cancer detection	High computational demand for 3D analysis
Xu et. al. [24]	2022	Chest CT scans	CNN (ISANET)	ISANET achieved 95.24% and 98.14% accuracy in classifying NSCLC subtypes	Advanced CNN architecture with attention mechanisms	Complexity in model design and training
Raju and Rao [25]	2022	Histopathological images	MobileNetV2, InceptionResnetV2	High classification accuracy up to 99.95%	Effective in identifying cancer tissues accurately, potential for automated	May not generalize well across different types of cancer images

JURNAL TEKNOINFO
Volume 18, Nomor 2, July 2024, Page 458-475
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

					systems	
Kalaivani et. al. [26]	2020	CT scans	DenseNet, AdaBoost	DenseNet and AdaBoost achieved 90.85% accuracy in lung cancer detection	Efficient image classification, focus on challenging instances	Data augmentation required for optimal performance
Agarwal et. al. [27]	2021	CT images	AlexNet-CNN	AlexNet-CNN achieved 96% accuracy in distinguishing malignant from benign tumors	Transfer learning with a focus on lung tumor classification	Reliant on transfer learning from existing models
Asuntha and Srinivasan [28]	2020	CT scans	FPSOCNN	Superior performance in accuracy, sensitivity, and specificity	Integration of advanced feature selection techniques	Requires complex optimization to achieve best results
Aluka et. al. [29]	2023	Lung cancer dataset	VGG-16, Inception V3, ResNet50	VGG-16 achieved 96% training and 93% validation accuracy in lung cancer detection.	Improved accuracy in lung cancer detection, contribution to medical imaging advancement	Some models did not perform as well
Pradhan et. al. [30]	2023	Patient health records	RNN (optimized by SA-SLnO)	PCA and t-SNE with SA-SLnO improved lung cancer diagnosis accuracy	Superior optimization algorithm, lower time complexity	Complex setup involving multiple optimization techniques
Khatun et. al. [31]	2023	Histopathology images	VGG19, ResNet50, EfficientNet B7, MobileNetV2	ResNet50 achieved 98% accuracy in classifying NSCLC subtypes	Effective in NSCLC subtype detection, promising avenue for refining models	Varying performance across different models
Nafea et. al. [32]	2023	Kaggle	EfficientNet B3	Achieved 96% accuracy in lung cancer detection	Increased detection accuracy	Requires substantial data augmentation
Shankara et. al. [33]	2023	LIDC	CNN	92.96% accuracy, 97.45% sensitivity, 86.08% specificity	High efficacy in diagnosis	Not specified
Prasanna et. al. [34]	2023	CT scans	CNN, SqueezeNet, MobileNet	Achieved 95.21% accuracy in binary classification	Utilization of lightweight networks	Limited discussion on limitations; future work on 3D analysis
Abid et. al. [35]	2023	LUNA-16, LIDC	CNN (DenseNet, AlexNet, VGG-16)	Achieved 98% accuracy, 98.93% sensitivity, 99% specificity	Effective early nodule identification	Requires large datasets for training
Kumar et. al. [36]	2023	-	UNet	Achieved 93.4% accuracy, 98.4% sensitivity, 97.1% specificity	Significantly enhanced lung cancer diagnosis	Focuses more on early stages; may need further validation
Biradar et. al. [37]	2022	Kaggle, Luna16	2D CNN	Achieved 88.76% accuracy in distinguishing malignant from benign nodules	Effective lung nodule edge extraction	Requires precise nodule detection enhancements
Anand et.	2023	Histopatho	CNN	Achieved 99.0% accuracy in cancer	Robust capability	Focus is broad,

JURNAL TEKNOINFO
Volume 18, Nomor 2, July 2024, Page 458-475
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

al. [38]		logy images		classification	in cancer classification	could be more specialized
Venkatesh and Bojja [39]	2022	CT images	CNN	Achieved approximately 98% accuracy in lung nodule classification	Integration of IoT for efficient data sharing	Requires integration of multiple complex systems
Malligeswari and Kavya [40]	2024	PET/CT scans	Res-U-Net	Achieved 96.45% accuracy in lung cancer classification	Classifying lung cancer into SCLC and NSCLC categories	Needs validation in clinical settings
Pandey and Bhandari [17]	2023	-	ML, DL	Highlighted AI's potential in improving accuracy and efficiency	ML and DL models for lung nodule segmentation and classification	More of a review than experimental; broad scope
Pandian et. al. [41]	2022	CT scan images	CNN, GoogleNet	Achieved 98% accuracy in lung cancer detection	Improved diagnostic processes and treatment planning	Limited to specific types of images; may need diverse datasets
Islam et. al. [42]	2023	Distinct datasets for COVID-19 and lung cancer	DCNN, GRU	Achieved 98.97% accuracy in lung cancer classification	Increased transparency in decision-making	Specific to certain conditions, not generalized
W. Abdul [43]	2020	LIDC, IDRI	CNN	Achieved 97.2% accuracy in lung cancer detection	Enhanced accuracy and prevention of overfitting	Needs further advancements for wider applications
Mothkur and Veerappa [44]	2022	CT images	Lightweight DNNs (2D CNN, SqueezeNet, MobileNet)	Achieved 85.21% accuracy in early-stage lung cancer detection	Low memory use in deep learning models	Potential limitations in accuracy compared to more robust models
Mathivanan et. al. [45]	2024	Multimodal data (imaging, genomics, clinical)	MFDNN	Achieved 92.5% accuracy in lung cancer classification	Integration of medical imaging, genomics, and clinical data	Requires stringent validation and regulatory compliance
Naseer et. al. [46]	2023	LUNA16	Modified U-Net, AlexNet, SVM	Achieved 97.98% accuracy in lung cancer detection	Precise lung cancer detection	Specific to one dataset; may not generalize
Malafaia et. al. [47]	2022	CT images	DL models (Various architectures)	Demonstrated stability and interpretability of DL models	Enhanced transparency and interpretability	Challenges with limited datasets
Wankhade and Vigneshwari [48]	2023	LIDC-IDRI	3D-CNN, RNN	High accuracy, sensitivity, and specificity in lung cancer stage identification	Enhancement of diagnostic precision	Requires validation in broader clinical settings
Lee et. al. [49]	2022	WSIs	CNN (VGG19,	Improved NSCLC detection accuracy with suitable tile sizes	Selecting suitable CNN	Focused on specific settings,

			InceptionResnetV2, EfficientNet b3)		architectures and tile sizes	may not be universally applicable
Shakeel et. al. [50]	2022	CT images	Deep neural networks, ensemble classifiers	Advanced DNNs and ensemble classifiers improved accuracy in detecting lung cancer from CT images	Overcoming image quality issues and utilizing sophisticated ML approaches	Complexity in the ensemble classifier setup

4. DISCUSSION AND COMPARISON

Using deep learning techniques has really improved how we detect lung cancer with medical imaging. This paper reviewed the application of various neural network architectures, including CNNs, RNNs, and GNNs, highlighting their effectiveness in enhancing diagnostic accuracy. These models can learn a lot from big datasets and tell us what type of lung cancer someone has accurately. Recent advances in deep learning for finding lung cancer have made a big difference in how accurate and fast we can diagnose it.

Dodia et al. [28] conducted a systematic review focusing on deep learning algorithms such as CNNs, U-Nets, and DenseNets applied to lung cancer detection from CT scans. They found that these methods, especially those tailored for medical imaging, show promising results in terms of accuracy, sensitivity, and specificity. However, challenges such as data imbalance and computational demands persist, indicating the need for further research. Similarly, Cao et al. [33] introduced E2EFP-MIL, an end-to-end weakly supervised method achieving high accuracy in classifying lung cancer subtypes from whole slide images (WSIs), demonstrating the potential of AI in improving diagnostic accuracy.

Additionally, Nahiduzzaman et al. [29] proposed a novel framework combining lightweight convolutional neural networks (LPDCNN) with a ridge extreme learning machine model for precise classification of lung cancer types using CT images, showcasing exceptional efficiency and accuracy. These advancements underscore the potential of deep learning in enhancing early and accurate lung cancer diagnosis, offering significant implications for patient management and treatment. Despite the progress, challenges such as data imbalance and computational demands persist, suggesting the need for further research to address these issues and explore multimodal data integration for more robust diagnostic models. Additionally, Shyamala Bharathi and Shalini [34] highlighted the significance of advanced deep learning models with attention mechanisms for improving lung cancer detection using CT and PET images, emphasizing the importance of efficient and interpretable models for early and accurate diagnosis. These studies show that deep learning could change how we diagnose lung cancer. It might lead to better, faster, and easier to understand models that could really help patients and healthcare workers.

5. CONCLUSION

This paper explains how deep learning, a type of advanced artificial intelligence, is transforming the way lung cancer is detected and treated. It uses complex tools like Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Graph Neural Networks (GNNs) to help medical experts diagnose lung cancer stages more quickly and accurately. This technological advancement significantly increases patients' chances of receiving effective treatment earlier, which improves their recovery outcomes. Overall, this progress represents a crucial development in the fight against lung cancer, offering new hope for better patient care. The study highlights the effectiveness of deep learning in accurately identifying cancer and distinguishing between different types of lung cancer. This capability is essential for developing customized treatments, allowing for therapies that are specifically designed for the exact type of lung cancer a patient is battling. Such targeted treatments represent a significant advancement in the field of personalized medicine, offering hope for more effective and tailored therapeutic approaches. However, even with these exciting developments, there are still big challenges to tackle. For example, these tools need more and varied data to learn from, which would make them better and stronger. Also, these systems need to be faster and cheaper to run. Using deep learning in hospitals means that tech experts and medical staff need to work together closely. This teamwork helps make sure that the technology is both helpful and safe for taking care of patients. Also, they must follow specific rules to keep patients safe and their information private. Working together like this helps bring new technology into hospitals smoothly and safely. Looking to the future, the paper suggests that future research should focus on using smart systems we already

JURNAL TEKNOINFO
Volume 18, Nomor 2, July 2024, Page 458-475
ISSN: 1693-0010(Print), ISSN: 2615-224X(Online)
Available online at https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index

have to help speed up learning about lung cancer. It also talks about combining different kinds of data to get a better understanding of the disease. This way, we might find new and better ways to treat and diagnose lung cancer. Exploring these ideas could really help make progress in fighting this illness. In summary, deep learning offers a powerful way to fight lung cancer, with the potential to spot and classify the disease early and accurately. Continued advances and teamwork in this area are key to fully unlocking AI's potential in healthcare, improving cancer treatment for patients everywhere.

REFERENCES

[1] S. Gupta et al., “Mailed fecal immunochemical test outreach for colorectal cancer screening: summary of a Centers for Disease Control and Prevention–sponsored summit,” CA. Cancer J. Clin., vol. 70, no. 4, pp. 283–298, 2020.
[2] P. Kaur, A. Singh, and I. Chana, “Computational techniques and tools for omics data analysis: state-of-the-art, challenges, and future directions,” Arch. Comput. Methods Eng., vol. 28, no. 7, pp. 4595–4631, 2021.
[3] A. Paul, “Adaptive tri-plateau limit tri-histogram equalization algorithm for digital image enhancement,” Vis. Comput., vol. 39, no. 1, pp. 297–318, 2023.
[4] A. Paul, P. Bhattacharya, and S. P. Maity, “Histogram modification in adaptive bi-histogram equalization for contrast enhancement on digital images,” Optik (Stuttg)., vol. 259, p. 168899, 2022.
[5] A. Paul, T. Sutradhar, P. Bhattacharya, and S. P. Maity, “Infrared images enhancement using fuzzy dissimilarity histogram equalization,” Optik (Stuttg)., vol. 247, p. 167887, 2021.
[6] S. Kumar and A. K. Bhandari, “Automatic tissue attenuation-based contrast enhancement of low-dynamic X-Ray images,” IEEE Trans. Radiat. Plasma Med. Sci., vol. 6, no. 5, pp. 574–582, 2021.
[7] Y. Kumar, S. Gupta, R. Singla, and Y.-C. Hu, “A systematic review of artificial intelligence techniques in cancer prediction and diagnosis,” Arch. Comput. Methods Eng., vol. 29, no. 4, pp. 2043–2070, 2022.
[8] S. K. Pandey, A. K. Bhandari, and H. Singh, “A transfer learning based deep learning model to diagnose covid-19 CT scan images,” Health Technol. (Berl)., vol. 12, no. 4, pp. 845–866, 2022.
[9] R. Kumar and A. K. Bhandari, “Spatial mutual information based detail preserving magnetic resonance image enhancement,” Comput. Biol. Med., vol. 146, p. 105644, 2022.
[10] S. Kumar, A. K. Bhandari, A. Raj, and K. Swaraj, “Triple clipped histogram-based medical image enhancement using spatial frequency,” IEEE Trans. Nanobioscience, vol. 20, no. 3, pp. 278–286, 2021.
[11] P. Kandhway, A. K. Bhandari, and A. Singh, “A novel reformed histogram equalization based medical image contrast enhancement using krill herd optimization,” Biomed. Signal Process. Control, vol. 56, p. 101677, 2020.
[12] S. K. Pandey and A. K. Bhandari, “A systematic review of modern approaches in healthcare systems for lung cancer detection and classification,” Arch. Comput. Methods Eng., vol. 30, no. 7, pp. 4359–4378, 2023.
[13] L. F. de F. Souza et al., “Internet of medical things: an effective and fully automatic IoT approach using deep learning and fine-tuning to lung CT segmentation,” Sensors, vol. 20, no. 23, p. 6711, 2020.
[14] L. Chen et al., “Multimodality Attention-Guided 3-D Detection of Nonsmall Cell Lung Cancer in 18 F-FDG PET/CT Images,” IEEE Trans. Radiat. Plasma Med. Sci., vol. 6, no. 4, pp. 421–432, 2021.
[15] D. Jalal and A. M. Abdulazeez, “A Review on Heart Disease Detection Classification Based on Deep Learning Algorithm,” Indones. J. Comput. Sci., vol. 13, no. 2, 2024.
[16] I. Ibrahim and A. Abdulazeez, “The role of machine learning algorithms for diagnosing diseases,” J. Appl. Sci. Technol. Trends, vol. 2, no. 01, pp. 10–19, 2021.
[17] S. Muawanah, U. Muzayanah, M. G. R. Pandin, M. D. S. Alam, and J. P. N. Trisnaningtyas, “Stress and Coping Strategies of Madrasah’s Teachers on Applying Distance Learning During COVID-19 Pandemic in Indonesia,” Qubahan Acad. J., vol. 3, no. 4, pp. 206–218, 2023, doi: 10.48161/Issn.2709-8206.
[18] L. Alzubaidi et al., Review of deep learning: concepts, CNN architectures, challenges, applications, future directions, vol. 8, no. 1. Springer International Publishing, 2021. doi: 10.1186/s40537-021-00444-8.
[19] M. M. Saeed, Z. Al Aghbari, and M. Alsharidah, “Big data clustering techniques based on spark: a literature review,” PeerJ Comput. Sci., vol. 6, p. e321, 2020.
[20] N. M. Farhan and B. Setiaji, “Indonesian Journal of Computer Science,” Indones. J. Comput. Sci., vol. 12, no. 2, pp. 284–301, 2023, [Online]. Available: http://ijcs.stmikindonesia.ac.id/ijcs/index.php/ijcs/article/view/3135
[21] W. Zhu, L. Xie, J. Han, and X. Guo, “The application of deep learning in cancer prognosis prediction,” Cancers (Basel)., vol. 12, no. 3, p. 603, 2020.
[22] A. Maharjan, “Machine Learning Approach for Predicting Cancer Using Gene Expression.” University of Nevada, Las Vegas, 2020.
[23] S. Babichev, I. Liakh, and I. Kalinina, “Applying a Recurrent Neural Network-Based Deep Learning Model for Gene Expression Data Classification,” Appl. Sci., vol. 13, no. 21, p. 11823, 2023.
[24] S. M. S. Abdullah and A. M. Abdulazeez, “Facial expression recognition based on deep learning convolution neural network: A review,” J. Soft Comput. Data Min., vol. 2, no. 1, pp. 53–65, 2021.
[25] S. Gupta and M. K. Gupta, “A comprehensive data‐level investigation of cancer diagnosis on imbalanced data,” Comput.

Intell., vol. 38, no. 1, pp. 156–186, 2022.
[26] X. Jing, Y. Zhou, and M. Shi, "Dynamic Graph Neural Network Learning for Temporal Omics Data Prediction," IEEE Access, vol. 10, pp. 116241–116252, 2022.
[27] X.-M. Zhang, L. Liang, L. Liu, and M.-J. Tang, "Graph neural networks and their current applications in bioinformatics," Front. Genet., vol. 12, p. 690049, 2021.
[28] X. Liu, K. W. Li, R. Yang, and L. S. Geng, "Review of Deep Learning Based Automatic Segmentation for Lung Cancer Radiotherapy," Front. Oncol., vol. 11, no. July, pp. 1–16, 2021, doi: 10.3389/fonc.2021.717039.
[29] S. Dodia, B. Annappa, and P. A. Mahesh, "Recent advancements in deep learning based lung cancer detection: A systematic review," Eng. Appl. Artif. Intell., vol. 116, no. August, p. 105490, 2022, doi: 10.1016/j.engappai.2022.105490.
[30] M. Nahiduzzaman, L. Faisal Abdulrazak, M. Arselene Ayari, A. Khandakar, and S. M. R. Islam, "A novel framework for lung cancer classification using lightweight convolutional neural networks and ridge extreme learning machine model with SHapley Additive exPlanations (SHAP)," Expert Syst. Appl., vol. 248, no. February, p. 123392, 2024, doi: 10.1016/j.eswa.2024.123392.
[31] S. K. Pandey and A. K. Bhandari, "A Systematic Review of Modern Approaches in Healthcare Systems for Lung Cancer Detection and Classification," Arch. Comput. Methods Eng., vol. 30, no. 7, pp. 4359–4378, 2023, doi: 10.1007/s11831-023-09940-x.
[32] D. Al-obidi and S. Kacmaz, "Facial Features Recognition Based on Their Shape and Color Using YOLOv8," in 2023 7th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), IEEE, 2023, pp. 1–6.
[33] M. Khalsan, M. Mu, E. S. Al-Shamery, L. Machado, M. O. Agyeman, and S. Ajit, "Intersection Three Feature Selection and Machine Learning Approaches for Cancer Classification," in 2023 International Conference on System Science and Engineering (ICSSE), IEEE, 2023, pp. 427–433.
[34] L. Cao et al., "E2EFP-MIL: End-to-end and high-generalizability weakly supervised deep convolutional network for lung cancer classification from whole slide image," Med. Image Anal., vol. 88, no. May, p. 102837, 2023, doi: 10.1016/j.media.2023.102837.
[35] P. Shyamala Bharathi and C. Shalini, "Advanced hybrid attention-based deep learning network with heuristic algorithm for adaptive CT and PET image fusion in lung cancer detection," Med. Eng. Phys., vol. 126, no. October 2023, p. 104138, 2024, doi: 10.1016/j.medengphy.2024.104138.
[36] L. N. Gumma, R. Thiruvengatanadhan, L. Kurakula, and T. Sivaprakasam, "A Survey on Convolutional Neural Network (Deep-Learning Technique) -Based Lung Cancer Detection," SN Comput. Sci., vol. 3, no. 1, pp. 1–7, 2022, doi: 10.1007/s42979-021-00887-z.
[37] Z. Xu, H. Ren, W. Zhou, and Z. Liu, "ISANET: Non-small cell lung cancer classification and detection based on CNN and attention mechanism," Biomed. Signal Process. Control, vol. 77, no. April, p. 103773, 2022, doi: 10.1016/j.bspc.2022.103773.
[38] M. S. N. Raju and B. S. Rao, "Classification of Colon and Lung Cancer through Analysis of Histopathology Images Using Deep Learning Models," Ing. des Syst. d'Information, vol. 27, no. 6, pp. 967–971, 2022, doi: 10.18280/isi.270613.
[39] N. Kalaivani, N. Manimaran, S. Sophia, and D. D. Devi, "Deep Learning Based Lung Cancer Detection and Classification," IOP Conf. Ser. Mater. Sci. Eng., vol. 994, no. 1, 2020, doi: 10.1088/1757-899X/994/1/012026.
[40] A. Agarwal, K. Patni, and D. Rajeswari, "Lung Cancer Detection and Classification Based on Alexnet CNN," Proc. 6th Int. Conf. Commun. Electron. Syst. ICCES 2021, pp. 1390–1397, 2021, doi: 10.1109/ICCES51350.2021.9489033.
[41] A. Asuntha and A. Srinivasan, "Deep learning for lung Cancer detection and classification," Multimed. Tools Appl., vol. 79, no. 11–12, pp. 7731–7762, 2020, doi: 10.1007/s11042-019-08394-3.
[42] M. Aluka, S. Ganesan, and V. P. Reddy, "International Journal of INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING A Comparative Study on Pre-Training Models of Deep Learning to Detect Lung Cancer," Orig. Res. Pap. Int. J. Intell. Syst. Appl. Eng. IJISAE, vol. 2023, no. 1, pp. 148–155, 2023, [Online]. Available: www.ijisae.org
[43] K. Pradhan, P. Chawla, and S. Rawat, "A deep learning-based approach for detection of lung cancer using self adaptive sea lion optimization algorithm (SA-SLnO)," J. Ambient Intell. Humaniz. Comput., vol. 14, no. 9, pp. 12933–12947, 2023, doi: 10.1007/s12652-022-04118-y.
[44] M. F. Khatun, M. R. Ajmain, and M. Assaduzzaman, "A Deep Learning Approach to Detect and Classification of Lung Cancer," 2023 Int. Conf. Adv. Technol. ICONAT 2023, pp. 1–6, 2023, doi: 10.1109/ICONAT57137.2023.10080801.
[45] A. A. Nafea, M. S. Ibrahim, M. M. Shwaysh, K. Abdul-Kadhim, H. R. Almamoori, and M. M. AL-Ani, "A Deep Learning Algorithm for Lung Cancer Detection Using EfficientNet-B3," Wasit J. Comput. Math. Sci., vol. 2, no. 4, pp. 68–76, 2023, doi: 10.31185/wjcms.209.
[46] C. Shankara, S. A. Hariprasad, and D. U. Latha, "Detection of Lung Cancer Using Convolution Neural Network," SN Comput. Sci., vol. 4, no. 3, pp. 1–8, 2023, doi: 10.1007/s42979-022-01630-y.
[47] K. R. Prasanna, R. V. Kumar, G. Mohanbabu, J. J. D. Selvam, J. Kumarnath, and S. P. Sellapandi, "Detection and Robust Classification of Lung Cancer Disease Using Hybrid Deep Learning Approach," 2023 Int. Conf. Data Sci. Agents Artif. Intell. ICDSAAI 2023, pp. 1–6, 2023, doi: 10.1109/ICDSAAI59313.2023.10452545.
[48] M. Abid, S. Akbar, S. Abid, S. A. Hassan, and S. Gull, "Detection of Lungs Cancer Through Computed Tomographic Images Using Deep Learning," 2023 4th Int. Conf. Adv. Comput. Sci. ICACS 2023 - Proc., pp. 1–6, 2023, doi: 10.1109/ICACS55311.2023.10089652.
[49] V. Kumar et al., "Improved UNet Deep Learning Model for Automatic Detection of Lung Cancer Nodules," Comput. Intell. Neurosci., vol. 2023, pp. 1–8, 2023, doi: 10.1155/2023/9739264.

[50] V. G. Biradar, P. K. Pareek, K. S. Vani, and P. Nagarathna, "Lung Cancer Detection and Classification using 2D Convolutional Neural Network," MysuruCon 2022 - 2022 IEEE 2nd Mysore Sub Sect. Int. Conf., pp. 1–5, 2022, doi: 10.1109/MysuruCon55714.2022.9972595.
[51] V. Anand, K. S. Gill, and S. Gupta, "Multi-class Classification of Colon and Lung Cancer using Deep Convolution Neural Network," Int. Conf. Sustain. Comput. Smart Syst. ICSCSS 2023 - Proc., no. Icscss, pp. 447–451, 2023, doi: 10.1109/ICSCSS57650.2023.10169254.
[52] C. Venkatesh and P. Bojja, "A Dynamic Optimization and Deep Learning Technique for Detection of Lung Cancer in CT Images and Data Access Through Internet of Things," Wirel. Pers. Commun., vol. 125, no. 3, pp. 2621–2646, 2022, doi: 10.1007/s11277-022-09676-0.
[53] N. Malligeswari and G. Kavya, "Classification of lung cancer with deep learning Res-U-Net and molecular imaging," Signal, Image Video Process., vol. 18, no. 1, pp. 325–333, 2024, doi: 10.1007/s11760-023-02635-7.
[54] R. Pandian, V. Vedanarayanan, D. N. S. Ravi Kumar, and R. Rajakumar, "Detection and classification of lung cancer using CNN and Google net," Meas. Sensors, vol. 24, no. September, p. 100588, 2022, doi: 10.1016/j.measen.2022.100588.
[55] M. K. Islam, M. M. Rahman, M. S. Ali, S. M. Mahim, and M. S. Miah, "Enhancing lung abnormalities detection and classification using a Deep Convolutional Neural Network and GRU with explainable AI: A promising approach for accurate diagnosis," Mach. Learn. with Appl., vol. 14, no. July, p. 100492, 2023, doi: 10.1016/j.mlwa.2023.100492.
[56] W. Abdul, "An Automatic Lung Cancer Detection and Classification (ALCDC) System Using Convolutional Neural Network," Proc. - Int. Conf. Dev. eSystems Eng. DeSE, vol. 2020-Decem, pp. 443–446, 2020, doi: 10.1109/DeSE51703.2020.9450778.
[57] R. Mothkur and B. N. Veerappa, "Classification of Lung Cancer Using Lightweight Deep Neural Networks," Procedia Comput. Sci., vol. 218, pp. 1869–1877, 2022, doi: 10.1016/j.procs.2023.01.164.
[58] S. S.K.B et al., "An enhanced multimodal fusion deep learning neural network for lung cancer classification," Syst. Soft Comput., vol. 6, no. October 2023, p. 200068, 2024, doi: 10.1016/j.sasc.2023.200068.
[59] I. Naseer, S. Akram, T. Masood, M. Rashid, and A. Jaffar, "Lung Cancer Classification Using Modified U-Net Based Lobe Segmentation and Nodule Detection," IEEE Access, vol. 11, no. April, pp. 60279–60291, 2023, doi: 10.1109/ACCESS.2023.3285821.
[60] M. Malafaia, F. Silva, I. Neves, T. Pereira, and H. P. Oliveira, "Robustness Analysis of Deep Learning-Based Lung Cancer Classification Using Explainable Methods," IEEE Access, vol. 10, no. September, pp. 112731–112741, 2022, doi: 10.1109/ACCESS.2022.3214824.
[61] S. Wankhade and V. S., "A novel hybrid deep learning method for early detection of lung cancer using neural networks," Healthc. Anal., vol. 3, no. April, p. 100195, 2023, doi: 10.1016/j.health.2023.100195.
[62] A. L. S. Lee, C. C. K. To, A. L. H. Lee, J. J. X. Li, and R. C. K. Chan, "Model architecture and tile size selection for convolutional neural network training for non-small cell lung cancer detection on whole slide images," Informatics Med. Unlocked, vol. 28, p. 100850, 2022, doi: 10.1016/j.imu.2022.100850.
[63] P. M. Shakeel, M. A. Burhanuddin, and M. I. Desa, "Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier," Neural Comput. Appl., vol. 34, no. 12, pp. 9579–9592, 2022, doi: 10.1007/s00521-020-04842-6.